Modeling Credit Risk through the Austrian Business Cycle: An Update of the OeNB Model

Michael Boss,
Gerhard Fenz,
Johannes Pann,
Claus Puhr,
Martin Schneider,
Eva Ubl[1]

In quantitative financial stability analysis, the link between the macroeconomic environment and credit risk is of particular importance when assessing the risk hidden in loan portfolios. Macroeconomic stress testing, in particular, which aims at measuring the impact of an economic crisis on individual banks or on the entire financial system, depends on means to quantitatively assess this link. Hence, the objective of this paper is to provide a methodological update of the OeNB's previous credit risk model that improves the capture of the relation between macroeconomic variables and probabilities of default for the main Austrian corporate sectors. In addition to the standard model based on individual macroeconomic variables, the paper explores solutions to two important challenges: first, the challenge related to the exploitation of potential information inherent in a larger macroeconomic data set and second, the problem that accounts for potential nonlinearity in the relation between credit and business cycles. The first issue is addressed via a regression model based on a principal components analysis that takes in a wider range of macroeconomic variables than commonly practiced. The second issue is addressed via a threshold approach. This paper presents the estimation results for the three different models and discusses them on the basis of an illustrative example.

JEL classification: C22, E32, G21
Keywords: Credit risk, credit cycle, financial stability, stress testing

1 Introduction

In recent years, a lot of effort has been put into modeling the link between macroeconomic variables and credit risk measures. Interest in this issue was first, driven by the new perspectives on risk based on the Basel II framework and has, more recently, been intensified by the financial crisis. As central banks and other supervisory authorities try to assess the impact of the financial crisis on the real economy and – once again – on banks' loan portfolios, understanding the relation of business and credit cycles has probably become more important than ever. This need has triggered a reassessment of commonly used approaches to measure credit risk with a focus on the capability of credit risk models to adequately capture downside risks, particularly in light of the ongoing crisis.

In terms of Basel II, the objectives of credit risk models are twofold: First, under Pillar I of the Internal Rating Based (IRB) Approach, banks can use their own credit risk forecasts as input for calculating regulatory capital. Second, banks are required to conduct stress tests under Pillar II. Forecasts as well as stress testing, however, not only matter for banks and their supervisors, but also for authorities concerned with financial market stability.

From a conceptual point of view, it should be possible to perform both forecasts and stress tests with a single model. But in practice, there are certain obstacles that have to be addressed. First of all, stress tests try to study the impact of shocks that are severe but plausible. However, such shocks are by definition hardly present over the sample horizon for which credit risk mod-

Refereed by:
Nicholas Vause and Joseph Noss, Bank of England,
Martin Summer, OeNB

[1] *Economic Research Division and Financial Stability Analysis and Surveillance Division. Corresponding author: eva.ubl@oenb.at. The authors would like to thank Thomas Scheiber and the OeNB's Foreign Research Division for technical support with the scenario, and the referees for helpful comments.*

els are usually estimated. Second, in pure forecasting exercises the ability to identify macroeconomic drivers of credit risk does not carry the same weight as in stress testing exercises, where storytelling is a fundamental feature. Moreover, Drehmann et al. (2006) argue that due to the presence of nonlinearity, standard econometric forecasting models such as linear vector autoregressions (VARs) are inadequate for stress testing.

All these arguments lead to the conclusion that time series models are inappropriate for stress tests, despite their superiority regarding the predictive power over the sample period. As an alternative, Drehmann et al. (2006) e.g. propose the application of a nonlinear methodology first published by Jordà (2005). The basic idea of Jordà's approach is to overcome the nonlinearity problem through estimating different approximation models (e.g. quadratic or cubic approximations) for each horizon of interest. Drehmann et al. (2006) emphasize that the results of their nonlinear VAR are significantly different to results obtained when using standard probit models.

Another strand of research focuses on the identification of threshold effects in credit risk stress testing. Gasha and Morales (2004) assess the impact of economic activity on nonperforming loans (NPLs) and conclude that advanced financial systems with low levels of NPLs appear to have an embedded self-correcting adjustment when NPLs exceed a minimum threshold. Bruche and González-Aguado (2007) propose another econometric approach which allows for time variation in default and recovery rate distributions via an unobserved Markov chain, which they interpret as the "credit cycle." One of the main conclusions of their empirical investigation is that the time variation in recovery rate distributions does amplify risk, but that this effect is much smaller than the contribution of the time variation in default probabilities to systematic risk.

Koopman et al. (2007) were presumably the first to tackle the problem of a certain degree of arbitrariness, choosing variables to take into account the numerous possibilities in modeling the link between macroeconomic variables and credit risk measures. They propose the application of a dynamic common factor model, as developed by Stock and Watson (2002), to overcome this problem. A related model using frequency domain analysis was implemented by Schneider and Spitzer (2004) to produce short-term forecasts of real Austrian GDP.

This paper is most closely related to the work of Boss (2002), on which the current OeNB model is based.[2] But there is other closely related literature, e.g. Virolainen (2004), Simons and Rolwes (2008) and Fiori et al. (2007), all of which make use of the framework linking the macro-environment to the business cycle, as originally proposed by Wilson (1997a and 1997b). Our contributions to the empirical credit risk literature are fourfold: First, we present the regression models for the Austrian corporate sectors. Second, we provide an illustrative example based on a macroeconomic scenario calculated with the OeNB's Austrian Quarterly Model (AQM).[3] This provides an

[2] *Although the methodological foundation of the OeNB model is to link macroeconomic variables to probabilities of default, the model described in Boss (2002) has been frequently updated and numerous improvements have been incorporated, most importantly the estimation of multiple models (one for each of the main Austrian corporate sectors).*

[3] *See section 5 for a detailed description.*

illustration and comparison of the performance of the different models. Third, in order to exploit the potential information inherent in a larger macroeconomic data set, we apply a principal component analysis (PCA) to a set of 24 Austrian macroeconomic variables.[4] This approach avoids the – usually arbitrary – selection of variables and makes use of the entire output of large-scale macroeconometric models such as the AQM. Fourth, in order to account for potential nonlinearity in the relation between credit and business cycles, we investigate a threshold approach.

The remainder of this paper is structured as follows: In section 2, we describe the underlying data set. Section 3 specifies the methodologies used and section 4 presents the results of the regression analysis. In section 5 we examine our models on the basis of a macroeconomic scenario to illustrate and discuss their dynamics. Finally, conclusions are drawn in section 6.

2 Data

When it comes to analyzing probabilities of defaults, we are fortunate in the sense that long historical time series are available for the Austrian economy. Our analysis is based on firm default frequencies for the period from 1970 to 2007. These default frequencies are calculated by dividing the number of quarterly defaults by the total number of firms in each sector; they are interpreted as sectoral default probabilities throughout the paper.[5] The number of firm defaults and the total number of firms were obtained from the Austrian creditor association Kreditschutzverband von 1870 and combined with additional information on the number of firms per sector from Statistics Austria.

For our analysis, the Austrian economy was divided into the following main sectors (with the number of firms at mid-2008 in parenthesis): agriculture (7,330), production and mining (22,912), construction (26,916), trade (56,224), tourism (22,723), transport (11,637), financial services (6,383), services (82,120), overall[6] (228,967).

In chart 1, the default probabilities of all sectors show an ascending trend at least for the 1970s. Most of the time series show evidence of structural breaks, in particular in the beginning of the 1990s. This is not surprising given the changes the Austrian economy underwent at the time, for example the privatization of large, formerly state-owned firms and the preparations for EU accession.

The macroeconomic variables were taken from the OeNB's macroeconomic database. Table 1 presents descriptive statistics of a representative sample of the Austrian macroeconomic variables included in our regressions. For a list of the 24 macroeconomic variables used for the PCA analysis, refer to table 8 in the appendix.

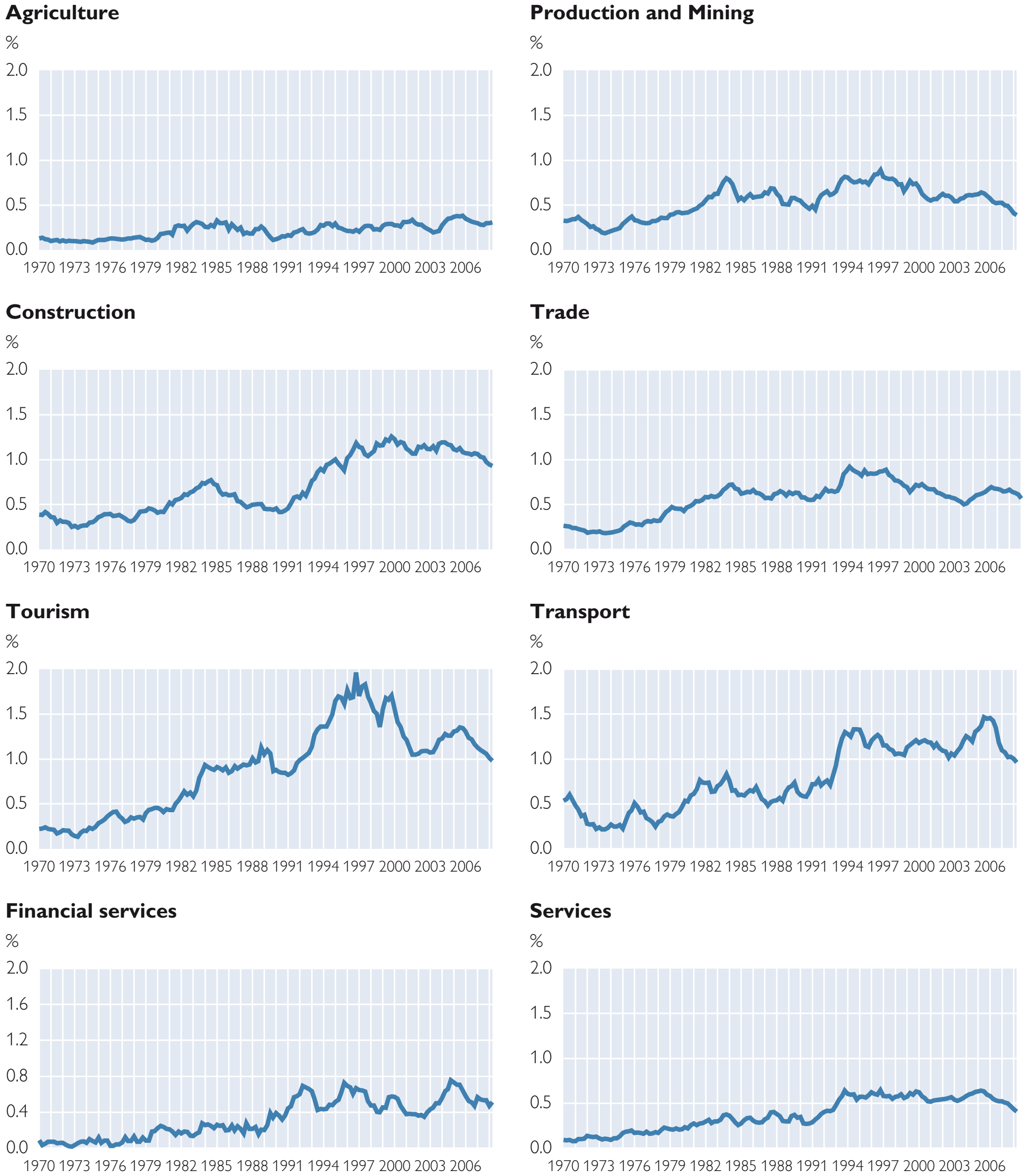
[4] *See table 8 in the appendix for a complete list of the 24 macroeconomic variables.*

[5] *Because of certain data limitations we use the moving average over four quarters in full knowledge of the problems of autocorrelation.*

[6] *"Overall" refers to the overall Austrian economy, excluding public services and the agricultural sector.*

Chart 1

Probabilities of Default for the Main Austrian Business Sectors (1970–2008)



Source: Kreditschutzverband von 1870, Statistics Austria.

3 Methodology

This section includes descriptions of the three models underlying our analysis (the standard regression model, the principal components analysis (PCA) and the threshold model) as well as the algorithm applied to select the optimal model.

3.1 Standard Regression Analysis

The average sectoral default probability at time t is modeled as a logistic function of an industry-specific "macroeconomic index" which, in turn, depends on the current values of the macroeconomic variables under observation. The initial logistic regression equation can be noted as:

Table 1

Descriptive Statistics of Quarterly Austrian Macroeconomic Variables

	Expected sign	Number of observations	Mean	Standard deviation
Cyclical indicators				
GDP, real (YER)	–	154	2,71	1,71
Industrial production, real (IPexE)[1]	–	154	3,57	4,49
Household indicators				
Private consumption, real (PCR)	–	154	2,59	1,98
PCR/GDP	–	158	0,57	0,02
Unemployment rate (URX)	+	154	2,83	1,44
Private sector disposable income, real (PYR)	–	154	2,74	2,41
Corporate indicators				
Average labor productivity (PRO)	–	154	2,25	1,61
Total investment, real (ITR)	–	154	2,53	4,80
Investment in equipment, real (IER)	–	154	3,16	6,56
IER/GDP	–	158	0,08	0,01
Unit labor costs (ULCN)	+	154	3,22	3,42
Storage (SCR)	+	152	–19,72	246,41
External indicators				
Exports, real (XTR)	–	154	6,31	4,30
XTR/GDP	–	158	0,34	0,12
Oil	+	154	15,84	51,42
USD	+	154	1,00	0,24
Price stability indicators and interest rates				
Consumer price index (CPI)	+	152	3,67	2,25
Short-term interest rate, real (STI real)	+	154	2,19	1,84
Long-term interest rate, real (LTI real)	+	154	3,36	1,56
Short-term interest rate, nominal (STI nominal)	+	154	5,85	2,41
Long-term interest rate, nominal (LTI nominal)	+	154	7,03	1,85

Source: OeNB.

[1] *Real industrial production excluding energy.*

Note: All variables are annual growth rates except for URX, USD, STI , LTI (real and nominal) and all ratios.

$$G(y_{t,s}) = p_{t,s} = \frac{1}{1+e^{-y_{t,s}}}$$

where $y_{t,s}$ denotes the industry-specific macroeconomic index at time t for sector s.[7]

We apply two different estimation methods for this equation. In the first estimation, we follow the approach proposed in Wilson (1997a and 1997b) and calculate "observed" values for the macroeconomic index y_t by simply taking the inverse of the logistic function based on the historically observed default probabilities:

$$y_t = -\ln\left(\frac{1}{p_t} - 1\right).$$

Since the macroeconomic index is not stationary, we conduct an ordinary least square regression for $\Delta y_t = y_t - y_{t-4}$.

This is reasonable for models with a long-term horizon such as ours (from 1970 to 2007), as such time series are subject to structural changes. Hence, an estimation of transformed levels could lead to wrong parameter estimates. The following regression equation was estimated:

$$\Delta\ y_t = \sum_{i=0}^{K} \beta\ \Delta x_{i,t} + \varepsilon_t = X_t \beta + \varepsilon_t \ \text{with}$$

$$\Delta x_{0,t} := 1$$

where y_t is the macroeconomic index, calculated according to the respective equation above. $\Delta x_{1,t}, \Delta x_{2,t}, \ldots, \Delta x_{K,t}$ denote

[7] *In the following, we skip the subindex s for reasons of simplicity as all sectors are modeled in the same way.*

the set of year-on-year (logarithmic or absolute) changes of the macroeconomic variables and $\beta_0, \beta_1, \beta_2, \ldots, \beta_K$ stand for the parameters to be estimated. They determine the direction and extent of the impact the factors have on the index and finally on the sector-specific default probability. These parameters are estimated by means of a linear regression, where the error term ε_t is assumed to be an independent, normally distributed random variable $\varepsilon_t \sim N(0,\sigma_\varepsilon)$.

Having calculated the coefficient vector $\hat{\beta}$, estimates for the default probabilities can be calculated on the basis of estimated changes of the macroeconomic index $\Delta\hat{y}_t$ as follows:

$\hat{p}_t = \dfrac{1}{1+e^{-(\hat{y}_{t-4}+\Delta\hat{y}_t)}}$, where $\hat{p}_t$ denotes the estimated probabilities of default.

Actually observed lagged values were used to calculate the first four estimations of the macroeconomic index.

The other method we apply for the initial logistic regression equation is based on the work of Papke and Wooldridge (1996), who estimate the default probabilities directly but, in contrast to common logistic regression, explicitly account for fractional data between 0 and 1. To account for this feature, we estimate default probabilities according to the following equations:

$p_t = G(\Delta X_t \beta) + \varepsilon_t$ and

$\varepsilon_t \sim N(0, \sigma_\varepsilon G(\Delta X\beta)\ \{1 - G(\Delta X\beta)\})$.

The estimation is done using a quasi-maximum likelihood method where the log likelihood is given by

$$\ln L(\beta) = \sum_{t=1}^{T} \{p_t \ln[G(\Delta X_t \beta)] + (1-p_t)\ln[\bar{G}(\Delta X_t \beta)]\},$$

where ΔX_t is the t-th row of ΔX.

This method uses an estimation technique superior to ordinary least squares, but which can be applied to levels of probabilities of default only. The problems related to the stationarity of the dependent variable could in this case be dealt with by including an AR(1) term. To get the maximum dependency on the macroeconomic variables and consequently a higher impact of the scenario on the probabilities of default, however, no AR terms were included in the models presented in this paper. Instead, a time variable was included to take into account the upward trend of the probabilities of default described in section 2.

3.2 Principal Components Analysis

Instead of estimating the probabilities of default by the changes of individual macroeconomic variables, we use a PCA and take the resulting factors as input for the regression analysis. A PCA is an orthogonal linear transformation that places the projection of the data with the greatest variance on the first coordinate. The other coordinates are chosen subsequently, so that they explain the maximum remaining variance subject to the condition of orthogonality. In this paper, the first five factors are taken into account and they explain 74% of the variance of the 24 variables.[8]

X is the $t \times n$ matrix of the standardized macroeconomic variables[9] of annual changes. We calculate the diagonal matrix of eigenvalues A and the matrix of eigenvectors V of $X'X$.

[8] *For a complete list of the 24 macroeconomic variables used in this analysis, refer to table 8 in the appendix.*

[9] *Macroeconomic variables are standardized by subtraction of the mean and division by the standard deviation.*

$X'XV = VA$
$V'X'XV = A$
$(XV)'(XV) = A$

The fraction of each eigenvalue λ_i over the sum of all eigenvectors is the variance explained by the i-th eigenvector. The first q eigenvectors sorted according to the size of the associated eigenvalues V_q constitute the orthogonal linear transformation of X described above. The reduced factors F are obtained by $F=XV_q$, with F being a $t \times q$ matrix. It is obvious from the equation that the factors must be orthogonal by definition. This property is advantageous in the context of a regression analysis as it helps avoid problems related to collinearity. However, PCA models might include variables that are not significant in explaining probabilities of default at all. In terms of stress testing, these variables might alter the results and hence interfere with a proper risk assessment.

3.3 Threshold Model

To assess the presence of potential asymmetries regarding the dependence of probabilities of default on the business cycle, we examine whether the estimated parameters – or variables included in the model selection process – of our standard model differ significantly from those of an exogenous threshold model. This could be an indication for an underestimation of risks of the standard modeling approach. We re-estimate our models under the following condition:

$$y_t = \delta_{1,t} \sum_{i=0}^{K} \beta_i \Delta x_{i,t} + \varepsilon_t +$$

$$+ \delta_{2,t} \sum_{i=0}^{K} \beta_i \Delta x_{i,t} + \varepsilon$$

$$\delta_{1,2,t} = \begin{cases} 1 \\ 0 \end{cases}$$

where y_t denotes the transformed probabilities of default, with $\delta_{1,t}=1$ for below-average growth of the Austrian economy of two consecutive quarters else $\delta_{1,t}=0$ and vice versa for $\delta_{2,t}$ which identifies the observations corresponding to above-average growth. The same model can be applied to Δy_t.[10]

3.4 Model Selection

To find the optimal multivariate model, we use the following model selection procedure. All macroeconomic variables under consideration are assigned to one of the following groups: cyclical indicators, price stability indicators, household indicators, corporate indicators, interest rates and external indicators. Then we estimate all possible models, including no more than one variable of each group per regression. The regression results are sorted by the value of the adjusted R-squared value for the logistic regression, respectively by the highest value for the quasi-likelihood estimator for the fractional logistic regression. The models with the wrong sign for the coefficients and with a t-value of below 1.2 are dropped. The same procedure is conducted for each sector. In a next step, the best model is selected from the sorted models, accounting also for other statistical properties such as AIC and BIC, F test and ML ratio.

[10] *This simple threshold model can be extended by including a threshold "kick-in" once growth has breached some low percentile or in periods where one variable (or a collection of variables) becomes highly volatile by historical standards.*

In the PCA analysis, an analogous procedure is applied with one group per factor. In addition to the five factors, the oil price as well as short- and long-term real interest rates are taken into account. For the PCA factors no sign restriction can be applied, which might lead to statistical artefacts.

4 Estimation Results

This section provides the estimation results for univariate regressions as well as for the three multivariate regressions outlined in the methodology section.

4.1 Results of a Univariate Analysis

In a first step we estimate all univariate[11] models with the fractional logistic regression model and with OLS regressions for the changes of the default probabilities for each sector. These estimations provide an indication of the dependency of the sectoral default probabilities on macroeconomic variables. T-values are documented in table 2.

It can be observed that GDP, private consumption, the unemployment rate and industrial production as well as the ratios of equipment investment to GDP and exports to GDP are significant under both regression methods for almost all economic sectors. By estimating the levels of the probabilities of default, significant coefficients can also be found for average labor productivity, private sector disposable income and real exports. By contrast, the regressions based on the changes of probabilities of default show significant estimates with the expected sign for real total investment, real equipment investment and the oil and consumer price index for most of the sectors. Among the cyclical indicators (GDP and industrial production), GDP has higher values for the fractional logistic regression models, while for the OLS models industrial production has a higher explanatory power for most of the sectors.

All PCA factors are highly significant in explaining the levels of probabilities of default in most of the business sectors. The first and second factors (which together explain 50% of the variance of all included macroeconomic variables) have a negative sign, therefore they move inversely to the probabilities of default. They are not significant for the agricultural sector, but that finding is consistent with economic intuition in case the first and second factors do indeed represent the business cycle.

[11] *Including a constant and, for the fractional logistic regression, an additional time variable.*

Table 2

t-Values of the Univariate Regression for Probabilities of Default and for the Changes of the Logit-Transformed Probabilities of Default

Dependent variable: Pt (Method: Fractional logistic regression)

Explanatory variable	Expected sign	Agriculture	Production	Construction	Trading	Tourism	Transport	Financial services	Services	Other
GDP, real (YER)	–	–4.20***	–3.72***	–3.53***	–3.00***	–1.63*	–1.96**	–1.16	–4.93**	–3.92***
Private consumption, real (PCR)	–	–3.28***	–1.50	–2.21**	–2.16**	–0.38	–0.57	0.66	–0.99	–1.6*
PCR/GDP	–	2.31	15.19	2.77	16.65	16.76	5.90	12.85	13.89	15.08
Unemployment rate (URX)	+	2.91***	3.75***	1.59*	2.83***	0.99	2.1**	3.04***	3.25***	2.69***
Average labor productivity (PRO)	–	–3.47***	–3.34***	–3.25***	–4.51***	–2.06**	–1.78*	–2.58**	–2.97***	–3.75***
Private sector disposable income, real (PYR)	–	–2.70**	–2.01**	–4.14***	–1.92***	–0.81	–2.00**	0.34	–4.13***	–2.5**
Total investment, real (ITR)	–	–0.62	0.14	0.21	0.03	2.26	0.06	0.26	–2.10**	0.26
Equipment investment, real (IER)	–	1.04	1.89	1.78	0.60	0.91	2.02	0.18	0.37	1.24
IER/GDP	–	–5.70***	–6.20***	–4.09***	–7.46***	–3.44***	–3.51***	–1.24***	–5.73***	–6.65***
Unit labor cost (ULCN)	+	–1.24	–3.81	–6.36	–1.73	–4.54	–3.62	0.12	–3.01	–3.86
Exports, real (XTR)	–	–1.36	–3.29***	0.13	–2.94***	–0.51	–0.23	–2.33**	–3.34***	–2.28**
XTR/GDP	–	–1.59	–23.40***	–7.30***	–20.29***	–27.93***	–7.03***	–14.08***	–21.55***	–24.91***
Short-term interest rate, real (STI real)	+	–0.45	–3.11	–4.53	–1.57	–1.15	–3.18	–1.72	–3.35	–2.85
Long-term interest rate, real (LTI real)	+	1.89*	0.61	–0.07	1.10	0.95	0.59	–1.33	0.74	0.79
Short-term interest rate, nominal (STI nominal)	+	–2.84	–6.57	–7.41	–4.19	–4.60	–4.78	–0.79	–6.24	–6.16
Long-term interest rate, nominal (LTI nominal)	+	–1.85	–5.03	–3.65	–2.99	–3.06	–2.50	–0.99	–4.06	–3.89
Industrial production, real (IPexE)[2]	–	–1.25	–2.57**	–1.73*	–1.97**	–0.40	–0.72	–1.54	–3.39***	–1.99**
Oil	+	–3.62	–4.23	–0.61	–3.95	–2.20	–1.05	–0.15	–1.94	–2.76
Consumer price index (CPI)	+	–1.36	–7.01	–3.18	–5.37	–9.47	–2.32	–1.96	–4.69	–6.27
Factor 1		–0.44	–4.06***	–5.58***	–1.90**	–3.3***	–2.77***	–3.20***	–4.58***	–3.82***
Factor 2		–2.68**	–4.27***	–4.38***	–4.14***	–3.03***	–4.82***	–2.81***	–5.70***	–4.88***
Factor 3		–2.28**	2.21**	2.02**	1.75*	1.27	0.96	3.15***	2.21**	1.84*
Factor 4		1.75*	–0.68	0.00	–0.94	–0.04	–0.58	–2.27**	–0.48	–0.39
Factor 5		–1.05	–2.58**	–3.58***	–1.7*	–1.25	–3.87***	–2.21**	–4.68***	–3.12***

Dependent variable: ΔYt (Method: OLS regression after logistic transformation)

Explanatory variable	Expected sign	Agriculture	Production	Construction	Trading	Tourism	Transport	Financial Services	Services	Other
GDP, real (YER)	–	–2.77***	–4.05***	–4.33***	–3.04***	–2.54***	–4.18***	1.25	–1.18	–4.61***
Private consumption, real (PCR)	–	–2.34**	–1.33	–3.33***	–0.59	–1.66**	–2.38**	1.08	0.17	–1.92
PCR/GDP	–	2.31	15.19	2.77	16.65	16.76	5.90	12.85	13.89	15.08
Unemployment rate (URX)	+	2.38**	7.81***	4.41***	4.25***	4.14***	2.88***	0.00	5.09***	5.89***
Average labor productivity (PRO)	–	–1.62	0.16	–1.20	–0.75	1.23	–3.18	0.89	0.84	–0.73
Private sector disposable income, real (PYR)	–	–2.29**	1.03	–2.94***	–0.18	–1.09	–1.10	1.42	0.84	–0.45
Total investment, real (ITR)	–	–2.86***	–4.96***	–6.42***	–4.42***	–4.84***	–4.87***	1.58	–0.68	–6.56***
Investment in equipment, real (IER)	–	–2.09**	–4.20***	–4.51***	–4.23***	–2.47**	–2.95***	0.66	–1.57	–4.57***
IER/GDP	–	–5.70***	–6.20***	–4.08***	–7.46***	–3.44***	–3.51***	–1.24	–5.73***	–6.66***
Unit labor costs (ULCN)	+	–1.63	0.16	–1.20	–0.75	1.23	–3.18	0.89	0.84	–0.73
Exports, real (XTR)	–	0.33	–5.44***	–2.26**	–2.65***	–0.95	–1.40	–0.99	–4.52***	–3.61***
XTR/GDP	–	–1.59	–23.39***	–7.30***	–20.30***	–27.93***	–7.03***	–14.08***	–21.55***	–24.91***
Short-term interest rate, real (STI real)	+	–0.28	–2.39	–3.69	–1.78	–4.08	–1.77	2.38	–3.20	–3.44
Long-term interest rate, real (LTI real)	+	0.94	2.05	–0.01	0.82	0.58	0.19	0.60	–0.46	0.87
Short-term interest rate, nominal (STI nominal)	+	–2.84	–6.57	–7.41	–4.19	–4.60	–4.78	–0.79	–6.24	–6.16
Long-term interest rate, nominal (LTI nominal)	+	–1.85	–5.03	–3.65	–2.99	–3.06	–2.50	–0.99	–4.06	–3.89
Industrial production, real (IPexE)[2]	–	–1.64*	–8.32***	–5.60***	–4.52***	–2.60***	–3.23***	–0.14	–5.78***	–6.50***
Oil	+	0.13	0.56	1.40	2.33**	2.88***	3.43***	2.33**	–0.26	2.25*
Consumer price index (CPI)	+	2.45**	1.81*	2.12**	3.92***	3.01***	1.14	0.45	2.33**	2.97***
Factor 1		0.68	–1.42	–1.07	0.57	–0.31	0.29	2.22**	–0.91	–0.55
Factor 2		0.64	–3.03***	–3.07***	–1.79*	–1.48	–1.85*	–1.39	–3.18***	–2.87***
Factor 3		2.91***	0.24	1.98*	–1.33	–1.03	–0.96	–0.81	–0.63	–0.41
Factor 4		–0.77	3.03***	0.33	0.40	3.15***	–0.03	–0.26	3.06***	2.46***
Factor 5		–0.92	–1.74	–3.66***	–1.16	–1.92	–4.05***	1.48	–2.09**	–2.76***

Source: OeNB.

[1] **, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively.*

[2] *Real industrial production excluding energy.*

Note: Data (1970–2007) include 155 observations per sector.

Table 3

Results of the Standard Regression Model (Method: OLS for the changes of the logistically transformed probabilities of default)

Sector: Overall
Time period: 1970–2008
Quarterly observations
Number of observations: 153

Dependent variable: ΔY_t OLS

Variable (lag)	Coefficient	t-Statistics	Probability
Constant	0.02	1.60	0.11
Industrial production	–1.30	–8.81	0.00
Unemployment rate	0.08	3.99	0.00
Unit labor costs (2)	2.63	4.89	0.00
Oil	0.10	5.75	0.00
Long-term interest rate, real	0.01	1.32	0.19
R-squared	0.55		
Adjusted R-squared	0.54		

Source: OeNB.

Table 4

Results of the Standard Regression Model (Method: Fractional Logistic Regression)

Sector: Overall
Time period: 1970–2008
Quarterly observations
Number of observations: 153

Dependent variable: P_t fractional logistic regression

Variable (lag)	Coefficient	t-Statistics	Probability
Constant	–3.78	–33.07	0.00
Time	3.79	42.04	0.00
Industrial production	–1.03	–7.38	0.00
Unemployment rate (4)	0.06	2.84	0.01
Investment in equipment/ GDP (4)	–18.45	–12.72	0.00
Exports/GDP (4)	–5.11	–27.56	0.00
Quasi-maximum likelihood	–5.31		
AIC	22.61		
BIC	40.64		

Source: OeNB.

4.2 Results of the Standard Regression Analysis

In a second step we estimate multivariate regression models for the differences of the transformed default probabilities as well as for their levels. We follow the model selection process described in subsection 3.4 and present the results of the two models in table 3 and 4.[12]

For the fractional logistic regression method we present two different models. The model presented in table 4 has the smallest quasi-maximum likelihood, the model presented in table 5 has a smaller AIC.

Depending on whether we estimate the levels or the changes of the logistically transformed probabilities of default, we include different variables in the selected models. The driving cyclical indicator for almost all models is industrial production, but for the fractional logistic model industrial production can be replaced by GDP without losing much of the model's explanatory power. Surprisingly, the macroeconomic variables in our models are very similar for almost all industrial sectors. For the fractional logistic models, the variables used besides industrial production are the unemployment rate, the investment in equipment-to-GDP ratio and the exports-to-GDP ratio. For the trade and tourism sectors, the nominal short-term interest rate was found to be significant, too. Variables for the services sector include the consumer price index, while for the transport sector they include oil instead of the exports-to-GDP ratio. For the models based on the changes of the macroeconomic index, industrial production, the unemployment rate, unit

[12] *Model results for all business sectors are available from the authors upon request.*

Table 5

Results of the Standard Regression Model (Method: Fractional Logistic Regression) without Ratios

Sector: Overall
Time period: 1970–2008
Quarterly observations
Number of observations: 153

Dependent variable: P_t fractional logistic regression

Variable (lag)	Coefficient	t-Statistics	Probability
Constant	−5.06	−67.25	0.00
Time	0.75	9.26	0.00
GDP	−6.78	−6.36	0.00
Unit labor costs	−12.37	−6.77	0.00
Exports (4)	−1.11	−2.91	0.00
Quasi-maximum likelihood	−5.32		
AIC	20.63		
BIC	35.65		

Source: OeNB.

Table 6

Results of the Regression Model Based on Principal Components

Sector: Overall
Time period: 1970–2008
Quarterly observations
Number of observations: 153

Dependent variable: P_t fractional logistic regression

Variable (lag)	Coefficient	t-Statistics	Probability
Constant	−5.15	−150.79	0.00
Time	0.49	7.44	0.00
Factor 1 (3)	−0.01	−2.02	0.05
Factor 2 (4)	−0.02	−3.63	0.00
Factor 3	0.03	6.55	0.00
Factor 4	0.11	13.00	0.00
Factor 5	−0.03	−3.43	0.00
Long-term interest rate, real	0.02	2.08	0.04
Oil (2)	0.25	9.12	0.00
Quasi-maximum likelihood	−3.70		
AIC	25.41		
BIC	47.50		

Source: OeNB.

labor costs or investment in equipment, oil and the short-term interest rate (real)[13] are significant in most of the sectors.

4.3 Results of the Principal Components Analysis

Moving on to the PCA, we re-estimate multivariate regression models for the levels of default probabilities, this time, however, including the five factors obtained from the PCA (lag 1 to 4). As not all 24 macroeconomic variables have been available since 1970, time series for the PCA start in 1986. This obvious disadvantage might be set off by the fact that structural breaks in the Austrian economy pose a lesser problem for the shorter time horizon. On this account and given the results of the univariate case, we only estimate the fractional logistic models for the PCA factors.[14] The inclusion of the PCA factors in the selection procedure is limited to one occurrence by factor in the final model. In addition, we include real interest rates as well as the oil price; in these cases, the maximum likelihood estimator and the expected sign are the selection criteria. As shown in table 5, all five factors, the oil price and the interest rates enter the selected model.

4.4 Results of the Threshold Model

Finally, we estimate the threshold model. Our intention is to verify that the data point segmentation of the threshold model is superior to an arbitrary selection when it comes to providing statistical proof for the as-

[13] *When including interest rates in the models, restrictions concerning the t-value must be very tolerant, coming to about 1.2, which corresponds to a p-value of about 0.2.*

[14] *Results of the OLS regression were not as promising as those of the fractional logistic regression.*

sumption that the model is consistent. However, we fail to verify this statement, which might be related to the fact that our threshold model approach is still very basic. Overall results do not look very promising, however. Further research to capture the nonlinearity of the business cycle will be necessary. For the downside part of the model, no cyclical indicators are significant; only exports and unit labor costs have statistically relevant explanatory power. For the upside part of the model, industrial production, unit labor costs and the oil price are significant variables.[15]

5 An Illustrative Example

To illustrate the applicability of our research, we analyze the impact of an adverse economic shock on probabilities of default according to the different models presented above.

5.1 Scenario Description

We assume a severe global recession which heavily impacts on the Austrian economy. This assumed downturn affects the Austrian economy mainly via three channels:

- a decline in demand for Austrian exports;
- a global reassessment of risk that drives up risk premiums on interest rates in Austria and thus causes a decline in investment and consumption. In addition, equity prices are assumed to fall in this scenario, exerting negative wealth effects on consumption.
- negative confidence effects, which amplify the negative wealth effects. Austrian households are assumed to step up their precautionary savings, and firms are assumed to postpone investment projects.

The impact of the downturn on the Austrian economy is simulated using the OeNB's Austrian Quarterly Model (AQM) (see Fenz and Spitzer, 2004, and Schneider and Leibrecht, 2006); it turns out to be severe (see table 6).

Box 1

The OeNB's Model for Quarterly Macroeconomic Analysis

The OeNB's model for quarterly macroeconomic analysis (Austrian Quarterly Model – AQM) is a small to medium-size macroeconomic model in the tradition of neoclassical synthesis. It is therefore in line with most models used by Eurosystem central banks. The long-term relationships are derived from a neoclassical optimization framework, whereas short-term dynamics are data driven. Adjustment to the real equilibrium is sluggish. Imperfections on goods and labor markets typically prevent the economy from adjusting instantaneously to the long-term equilibrium. In the current version of the AQM, the formation of expectations is strictly backward looking. The relatively small scale of the model keeps the structure simple enough for projection and simulation purposes, while incorporating a sufficiently detailed structure to capture the main characteristics of the Austrian economy. The main behavioral equations are estimated using the two-step Engle-Granger technique. The model currently consists of 146 variables.

[15] *Result tables are available from the authors upon request.*

Table 7

GDP Growth According to the Assumed Scenario

	Year 1				Year 2				Year 3			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
	Year-on-year change in %											
Austrian GDP growth	1.9				–1.6				–0.4			
	Quarter-on-quarter change in %											
Austrian GDP growth	0.6	0.4	0.1	–0.1	–0.9	–0.7	–0.5	–0.3	–0.1	0.1	0.3	0.5

Source: OeNB.

In this three-year scenario, the GDP growth rate for Austria in quarter-on-quarter terms turns negative at the end of the first year and remains negative for six consecutive quarters. The trough is reached in the first quarter of the second year with a quarterly decline of GDP of –0.9%. In the third year of our scenario, GDP growth turns positive, but remains below potential growth until the end of the scenario horizon. Such a long economic downturn is an extraordinary event, which in reality has not been observed in Austria since World War II. The slump in economic activity in our scenario is mainly caused by a decline in exports and business investment, while the negative impact on private and public consumption is significantly smaller.

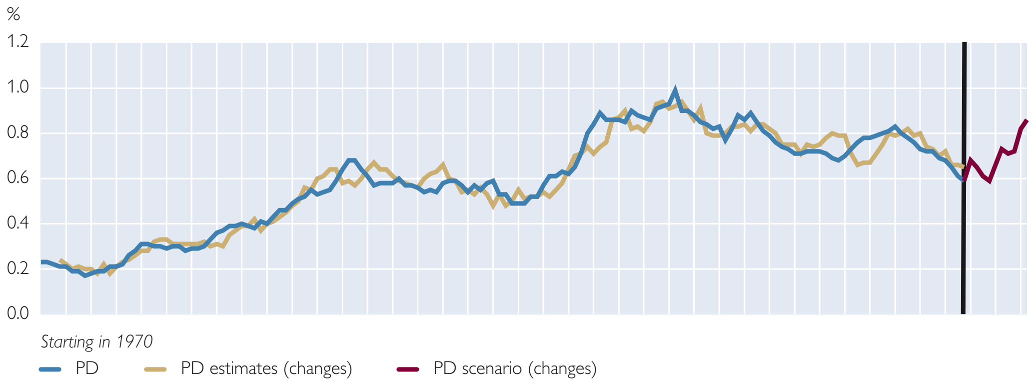
Given the scenario described above, different sensitivities of the probabilities of default can be observed that depend on (1) the model as well as (2) the sector. In the following two subsections we present the impact of the scenario on the overall sector according to the standard regression and the PCA models presented in section 3.[16]

5.2 Impact of the Scenario Based on Standard Regression Analysis

The graph presented in chart 2 is based on the regression estimations presented in table 3. It shows the impact of the as-

Chart 2

Probabilities of Default (PD) for the Overall Sector: OLS Regression Graph and Scenario Impact

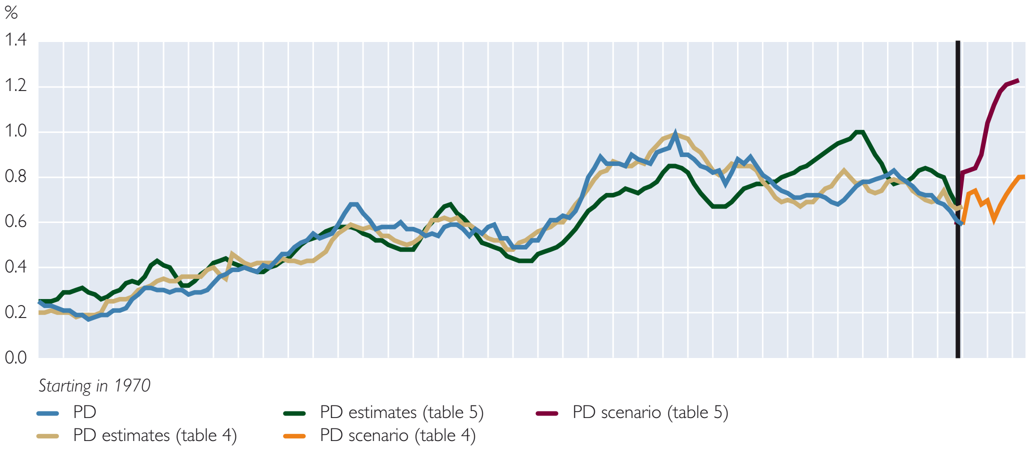


Source: OeNB.

[16] *Graphs and impact tables for all other sectors are available from the authors upon request.*

Chart 3

Probabilities of Default (PD) for the Overall Sector: Fractional Logistic Regression Graph and Scenario Impact



Source: OeNB.

sumed scenario according to standard regressions based on the differences of transformed default probabilities. This model has a rather small impact under the scenario. The changes of the probabilities of default depend on a lag of four quarters; therefore the impact of the variables is time delayed and shows a seasonal trend.

Chart 3 shows the impact of the fractional regression models presented in table 4 and 5. The graph based on the estimation presented in table 4 has a better fit than the graph based on table 5 results, but in the scenario the probabilities of default increase only by about 30%, which is quite similar to the model based on changes. The smaller impact of the scenario on the probabilities of defaults for the model based on table 4 results, which includes ratios of macroeconomic factors as explanatory variables, might be due to the fact that in the scenario in which both macroeconomic variables move in the same direction the change in the ratios is smaller than the change in the macroeconomic variable itself.

The impact of the scenario on the model based on table 5 shows a 100% increase in the probabilities of default. Independent of econometric arguments, from a supervisory perspective we feel more comfortable using this model because of its quicker response and the more pronounced increase in probabilities of default, and because it provides more conservative estimates of stress impact.

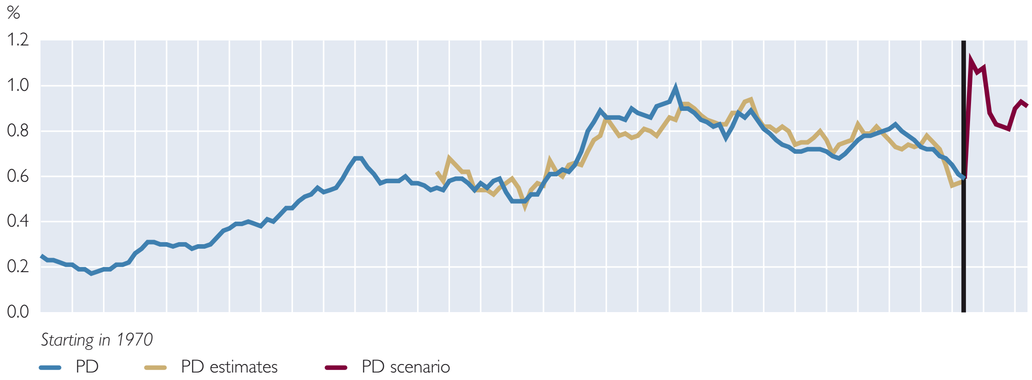
5.3 Impact of the Scenario Based on the Principal Components Analysis

The graph in chart 4 is based on the PCA model; regression estimations are the same as presented in table 6.[17] We observe a good fit and an increase of the probabilities of default of about 100% in the scenario. However, the probabilities of default show a considerable increase in the first quarter of the scenario and a drop in subsequent quarters. This property could not be ob-

[17] *Since no broad macroeconomic dataset is available for the period before 1987, the estimation period of the PCA model is shorter than the periods considered in the other models.*

Chart 4

Probabilities of Default (PD): PCA Fractional Logistic Regression Graphs and Scenario Impact



Source: OeNB.

served at any point in time over the horizon of our input time series. Further research is called for to analyze this surprising result.

6 Conclusion

With this paper we aim at achieving the methodological improvements necessary to bridge the gap between macroeconomic forecasting and credit risk modeling in order to run consistent macroeconomic stress tests. The ongoing crisis, in particular, highlights the need to quantitatively assess the impact of a possible economic deterioration on individual banks' loan portfolios or even on the entire financial system. In a period of systemic fragility it is of utmost importance to have a clear view of potential future credit defaults, as policymakers are more than ever called upon to help draft the appropriate policy response.

In that light, the objective of this paper – i.e. improving the OeNB models that link Austrian default probabilities to macroeconomic variables – is as timely as it could be. Next to standard regression models, we explore a PCA of 24 macroeconomic variables and an external threshold model. The models based on factors derived from the PCA are statistically significant and show a good fit. An economic interpretation of these factors, and hence the story-telling capacity based on this model, however, suffer from the methodology's lack of transparency and tractability. Moreover, results from our illustrative example were quite surprising. Under the assumed scenario, the probabilities of default increased rapidly in the first quarter but decrease later on. These results are not in line with economic intuition. For the threshold approach, no consistent models were found, as simulations based on arbitrary data point segmentation suggest that the models were driven by statistical artefacts. However, including the nonlinearity of the business cycle might increase the value of our threshold model and will be subject to further research at the OeNB.

As our attempt to address two of the main shortcomings in modeling the link between credit and business cycles – namely arbitrary variable selection and nonlinearity – has yielded no convincing results so far, we returned to more traditional modeling approaches. Two different methods to estimate standard regression are presented in this paper, each with its individual advantages and disadvantages. Regressions based on the changes of the transformed probabilities of defaults avoid problems like the nonstationarity of the default probabilities and other concerns related to structural breaks in economic time series. Fractional logistic regression, however, is a superior estimation method to OLS and is especially suited for data between 0 and 1.

Moreover, models for the changes of the logarithmically transformed probabilities of default react rather weakly to the scenario of our illustrative example. The increase in probabilities of default was comparatively small for such a severe scenario. In our example, models based on fractional logistic regression show a higher sensitivity than other models. In fact, probabilities of default double under the presented scenario. Surprisingly, in almost all corporate sectors, similar macroeconomic variables prove to be significant, but they differ depending on whether we estimate on the basis of changes in, or on the basis of levels of, the probabilities of default. In the light of this observation, it is even more surprising that our results show that the different models have comparable explanatory power, while at the same time showing vastly different properties regarding their reaction to the economic scenario assumed in our illustrative example.

By way of conclusion we can say that from the supervisory point of view, we prefer using the models based on fractional logistic regression, as they provide conservative estimates of probabilities of default in times of economic distress. Given the importance of the topic further research is called for, however, to support the continuous improvement of the models used to forecast and stress probabilities of default and, hence, of our capacity to properly assess the impact of the macroeconomic environment on credit risk.

References

Bruche, M. and C. González-Aguado. 2007. Recovery Rates, Default Probabilities and the Credit Cycle. Working Paper. Available at SSRN: http://ssrn.com/abstract=934348

Boss, M. 2002. A Macroeconomic Credit Risk Model for Stress Testing the Austrian Credit Portfolio. Financial Stability Report 4. OeNB. 64–82.

Drehmann, M., A. J. Patton and S. Sorensen. 2006. Corporate Defaults and Macroeconomic Shocks: Non-linearities and Uncertainty. Working Paper. Available at http://www.economics.ox.ac.uk/members/andrew.patton/research.html

Fenz, G. and M. Spitzer. 2004. AQM: The Austrian Quarterly Model of the Oesterreichische Nationalbank. Macroeconomic Models and Forecasts for Austria. Workshops No. 5. OeNB. 11–60.

Fiori, R., A. Foglia and S. Iannotti. 2007. Estimating Macroeconomic Credit Risk and Sectoral Default Rate Correlations for the Italian Economy. Working Paper. Bank of Italy.

Gasha, G. J. and A. R. Morales. 2004. Identifying Threshold Effects in Credit Risk Stress Testing. IMF Working Paper 04/150.

Jolliffe, I. T. 2002. Principal Component Analysis. Springer Series in Statistics. 2nd edition. New York: Springer.

Jordà, O. 2005. Estimation and Interference of Impulse Responses by Local Projections. American Economic Review 95(1). 119–147.

Koopman, S. J., A. Lucas and B. Schwaab. 2007. Forecasting Cross-Sections of Frailty-Correlated Default. Working Paper. Available at SSRN: http://ssrn.com/abstract=1113047

Papke, L. and J. Wooldrigde. 1996. Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. Journal of Applied Econometrics 11. 619–663.

Schneider, M. and M. Leibrecht. 2006. AQM-06: The Macroeconomic Model of the OeNB. Working Paper No. 132. OeNB.

Schneider, M. and M. Spitzer. 2004. Forecasting Austrian GDP Using the Generalized Dynamic Factor Model. Working Paper 89. OeNB.

Simons, D. and F. Rolwes. 2008. Macroeconomic Default Modelling and Stress Testing. Presentation held at the workshop "Stress Testing of Credit Risk Portfolios: The Link between Macro and Micro", hosted by the Basel Committee on Banking Supervision and De Nederlandsche Bank. 7 March 2008. http://www.bis.org/bcbs/events/rtf08simonspres.pdf

Stock, J. and M. Watson. 2002. Forecasting Using Principal Components from a Large Number of Predictors. Journal of Money, Credit and Banking 38(5). 1211–1261.

Virolainen, K. 2004. Macro Stress Testing with Macroeconomic Credit Risk Model for Finland. Paper prepared for the ECB Workshop of Financial Stability, 16-17 June 2004.

Wilson, T. 1997a. Portfolio Credit Risk (I), Risk 10 (9). 111–116.

Wilson, T. 1997b. Portfolio Credit Risk (II), Risk 10 (10). 56–61.

Appendix

Table 8

Macroeconomic Variables Transformed in the Principal Component Analysis (PCA)

Variables included in the PCA	Number of observations	Mean	Standard deviation
Total capital cost (CAC)	84	5.43	2.79
Private credit, amount outstanding (CPN)	84	6.46	2.67
Domestic demand, real (DDR)	84	2.22	1.29
Government budget balance (GB)	84	279	1317
Government debt gross (GDN)	84	5.20	4.12
Government disposal income (GYN)	84	4.27	3.97
Harmonised index of consumer prices (HICP)	84	1.94	0.83
Interest payments on government debt (INN)	84	3.94	6.26
Total investment, real (ITR)	84	2.49	2.37
Real marginal product of capital (MPC)	84	0.03	0.00
Imports, real (MTR)	84	5.71	3.83
Net foreign assets (NFA)	84	3.04	20.70
Net factor income (NFN)	84	34	74
Private consumption, real (PCR)	84	2.24	1.46
Direct tax paid by households (PDN)	84	4.82	5.43
Average labor productivity (PRO)	84	1.83	0.72
Private sector disposal income, real (PYR)	84	2.47	2.08
Total tax revenues (TOTREV)	84	3.60	3.77
Unit labor costs, adjusted (ULA)	84	0.63	0.04
Unemployment rate (URX)	84	3.97	0.60
Value added tax (VAT)	84	3.29	3.29
Real compensation per employee (WURYD)	84	0.95	0.84
Export, real (XTR)	84	6.20	3.49
GDP, real (YER)	84	2.48	1.07

Source: OeNB.

Note: All values are annual growth rates, except MPC, INN, ULA and INN levels.